



SEMPR: **The TSP Software Engineering** **Measured Performance** **Repository**

TSP Symposium, Pittsburgh
Nov 4, 2014

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA

William Nichols, Yasutaka Shirai

Report Documentation Page			<i>Form Approved OMB No. 0704-0188</i>	
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>				
1. REPORT DATE 04 NOV 2014	2. REPORT TYPE	3. DATES COVERED 00-00-2014 to 00-00-2014		
4. TITLE AND SUBTITLE SEMPR: The TSP Software Engineering Measured Performance Repository			5a. CONTRACT NUMBER	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)			5d. PROJECT NUMBER	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Carnege Mellon University, Software Engineering Institute, Pittsburgh, PA, 15213			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited				
13. SUPPLEMENTARY NOTES Team Software Process (TSP-2014) Symposium, 3-6 Nov, Pittsburgh, PA.				
14. ABSTRACT				
15. SUBJECT TERMS				
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 27
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified		

Document Markings

Copyright 2014 Carnegie Mellon University

This material is based upon work funded and supported by TSP cost recovery and TSP partner and licensing fees. under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center sponsored by the United States Department of Defense.

References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by Carnegie Mellon University or its Software Engineering Institute.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

This material has been approved for public release and unlimited distribution except as restricted below.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

Team Software ProcessSM and TSPSM are service marks of Carnegie Mellon University.

DM-0001838



Software Engineering Institute | Carnegie Mellon University

TSP Symposium 2014
William Nichols Nov 4, 2014
© 2014 Carnegie Mellon University

Agenda

1. Introduction
2. SEMPR data and analysis
3. Conclusion



Agenda

1. Introduction
2. SEMPR data analysis
3. Conclusion



Purpose of this presentation

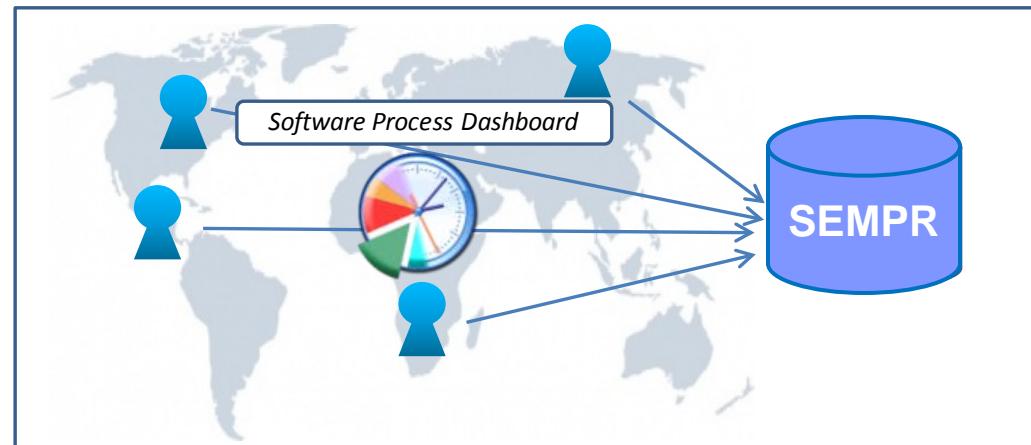
- This presentation tells...
- Project overview in SEMPR
- Benchmark planning parameters in SEMPR
- Benchmark project level performance and work item (component) level performance



About SEMPR

- Software Engineering Measured and Performance Repository

- SEI has collected data from organizations that have adopted TSP in SEMPR
- Stores project data in Tuma Solutions Team Process Data Warehouse
 - From 109 project cycles (in this report)
 - Used the Software Process Dashboard



How did we measure data quality in SEMPR

- Time log and defect log have high correctness and consistency by automatic data recording.
- Size log and task log have low correctness by manual data recording.

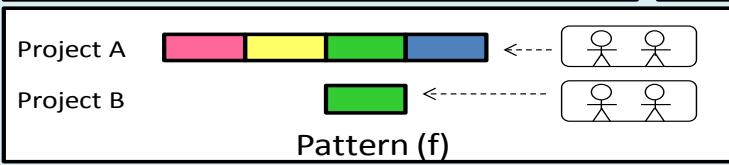
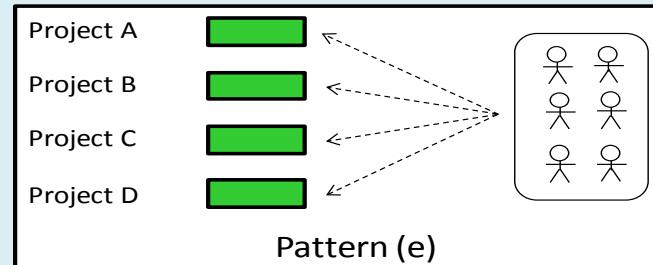
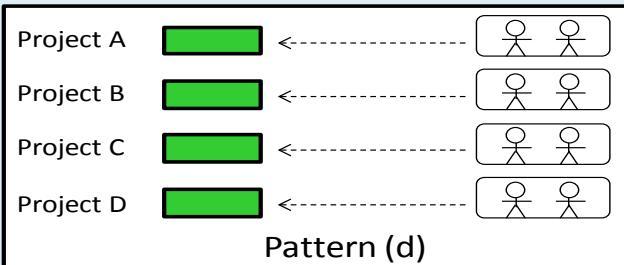
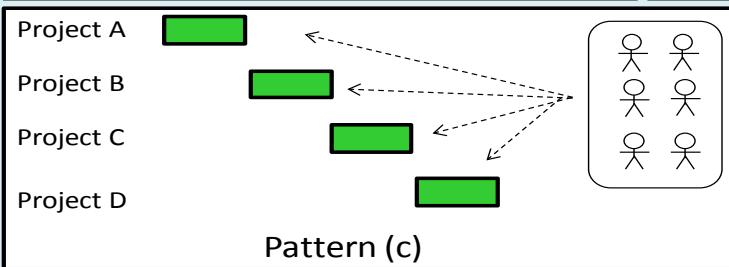
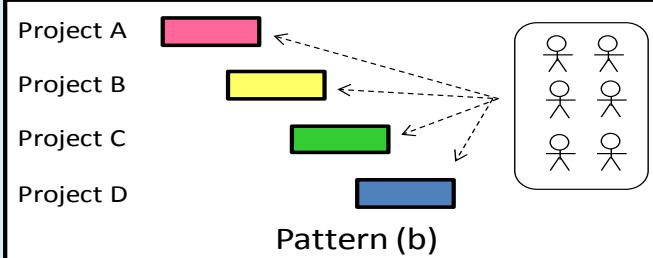
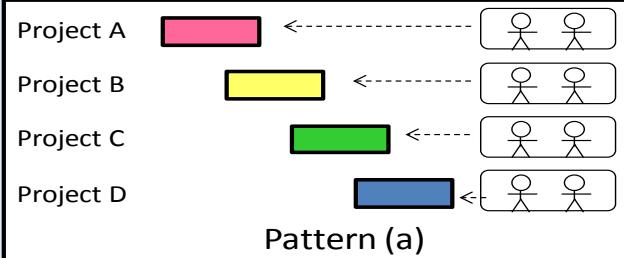


What do the data tell us?

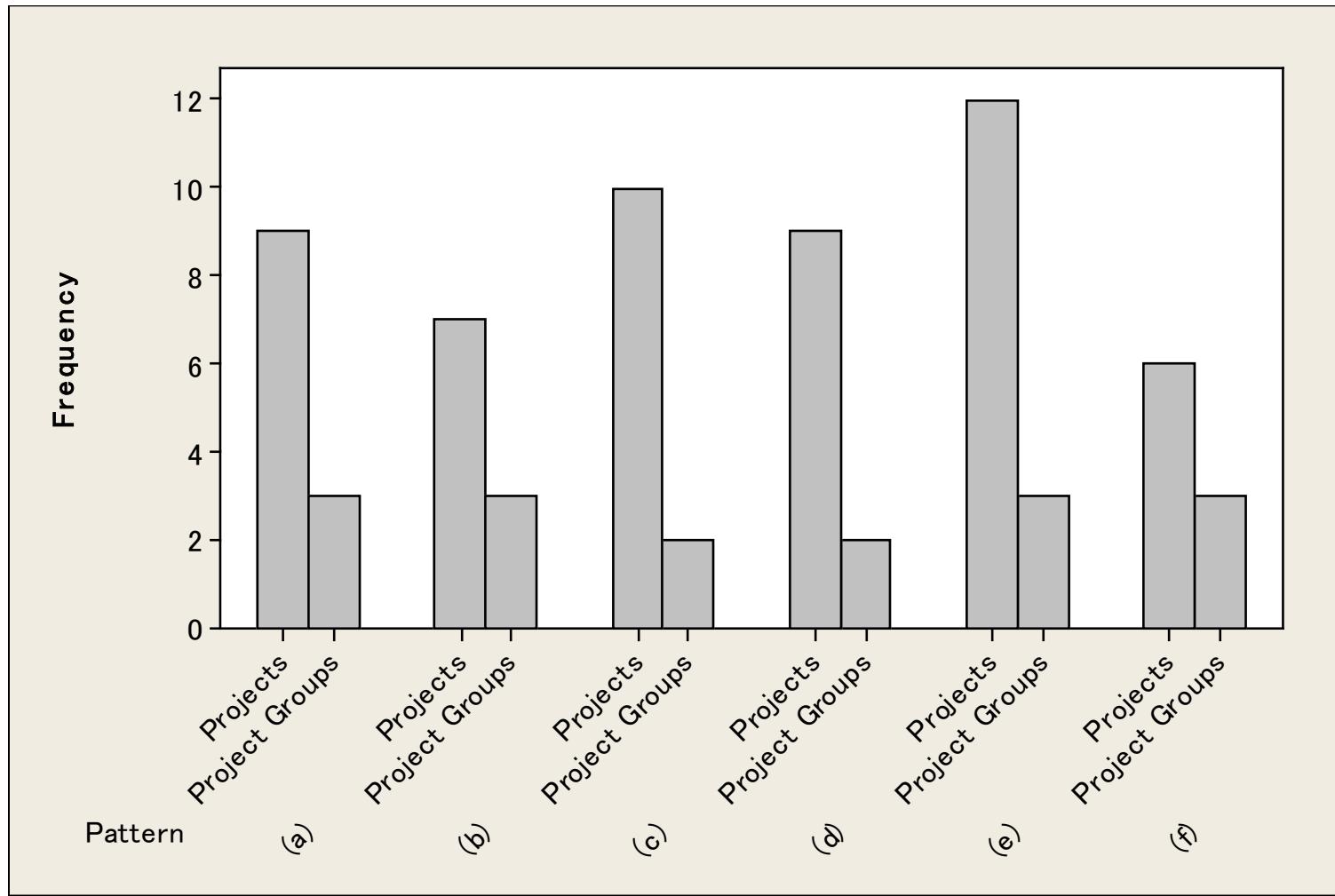
1. Introduction
2. SEMPR data analysis
3. Conclusion



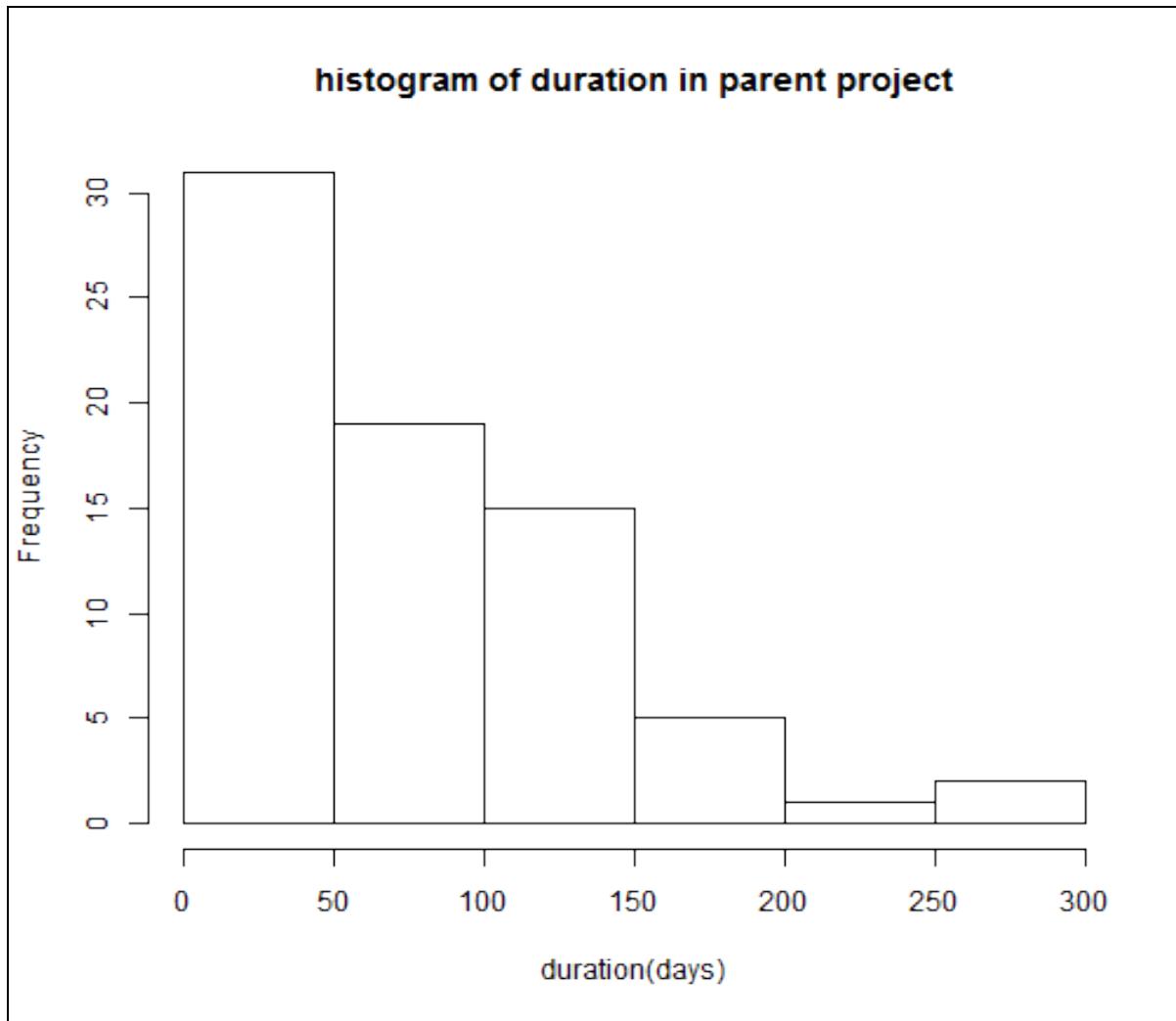
How are projects organized?



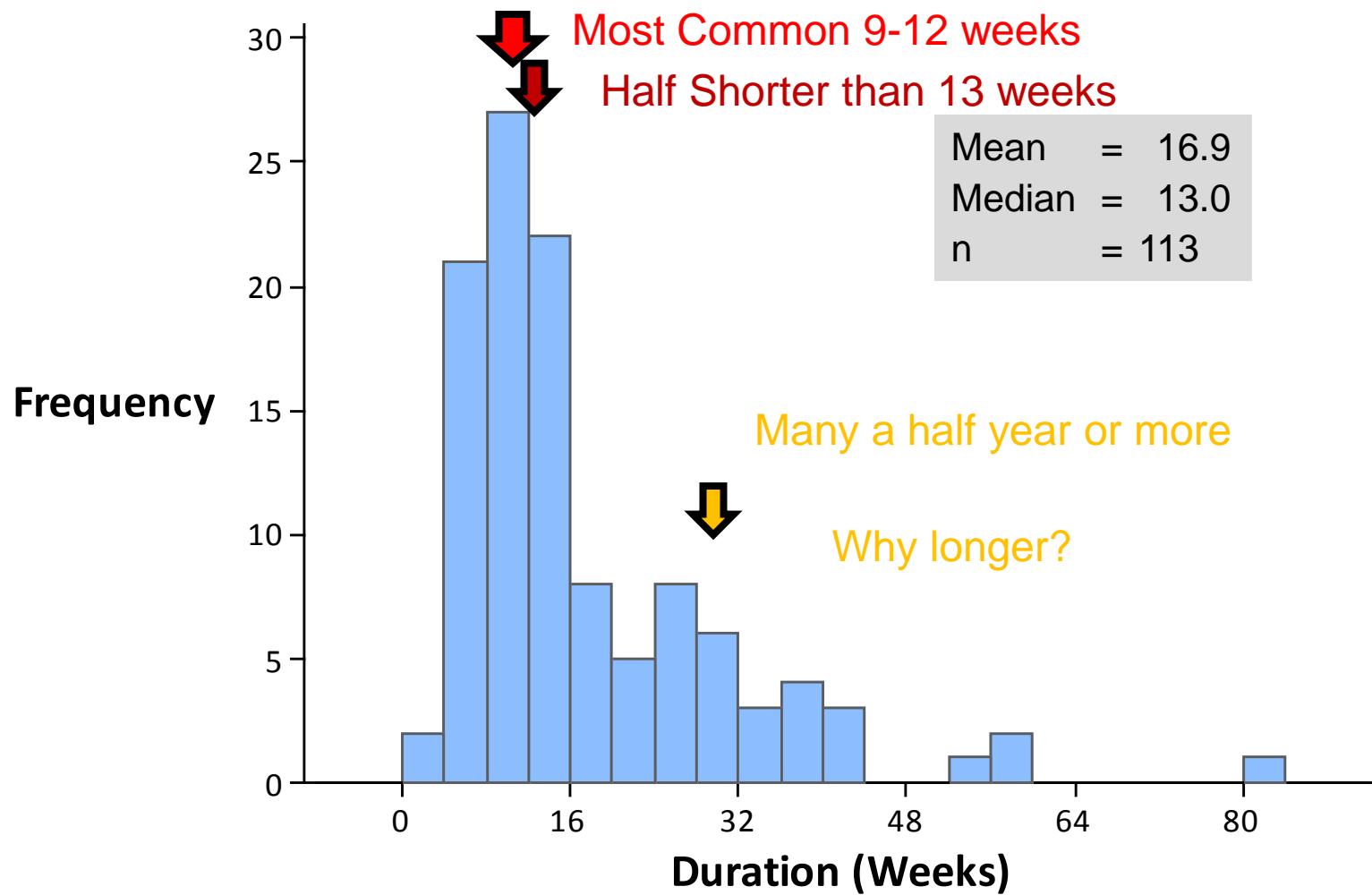
How many projects are found in each pattern?



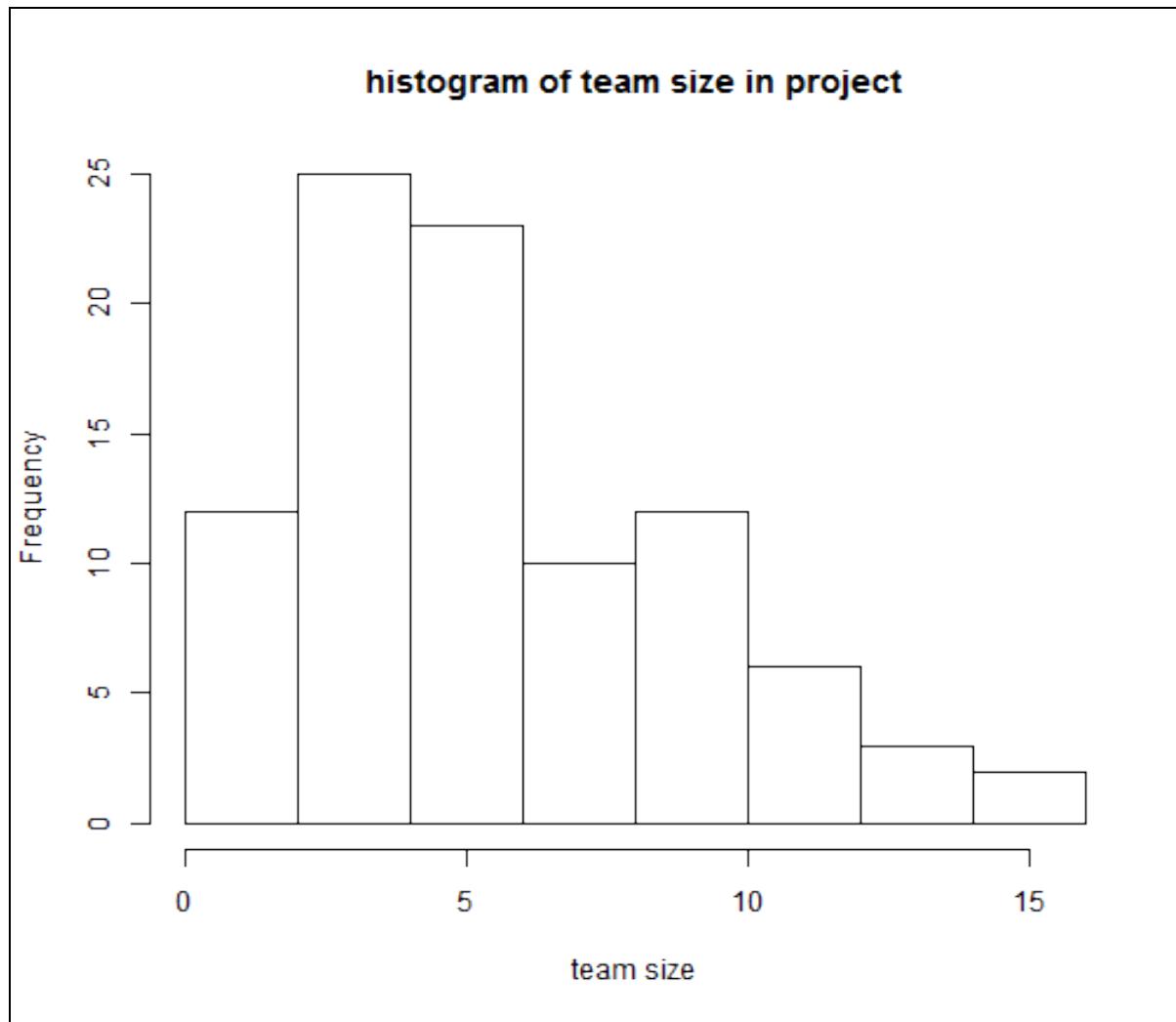
What were the project durations?



What are the planning period durations?

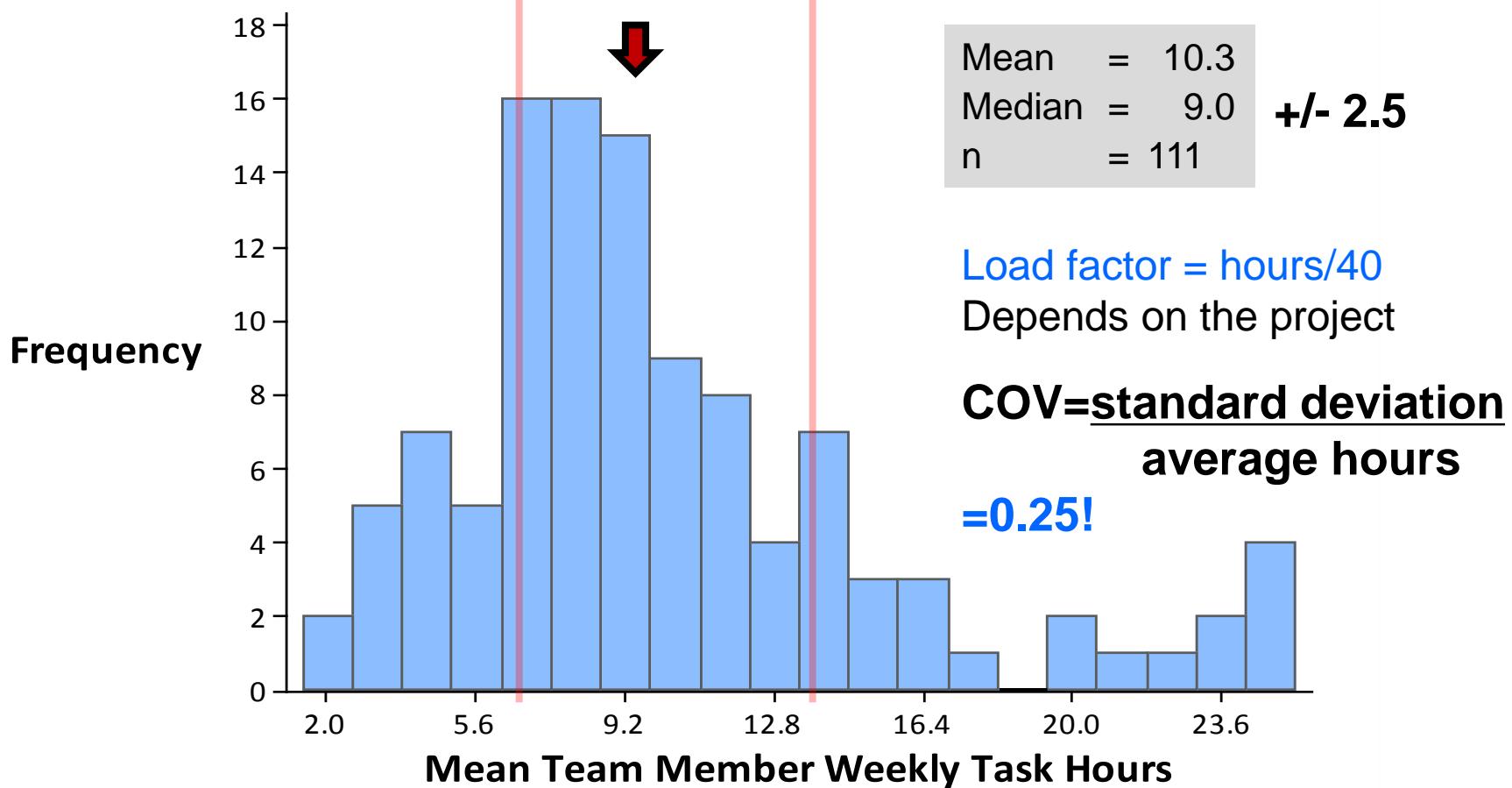


How many team members on projects?



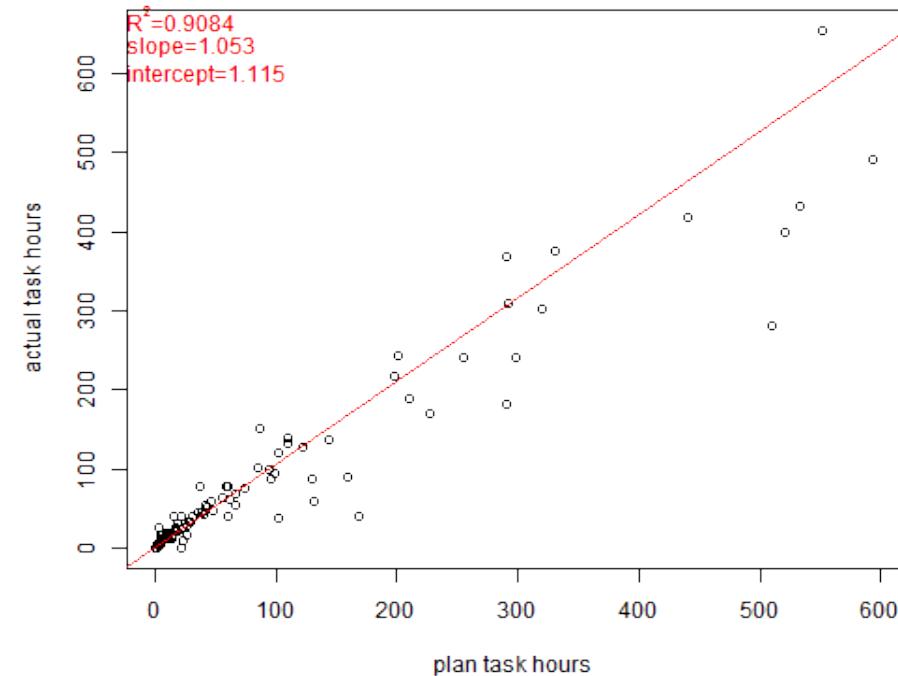
How many task hours per week?

mean Team Member Weekly Direct Hours per Week



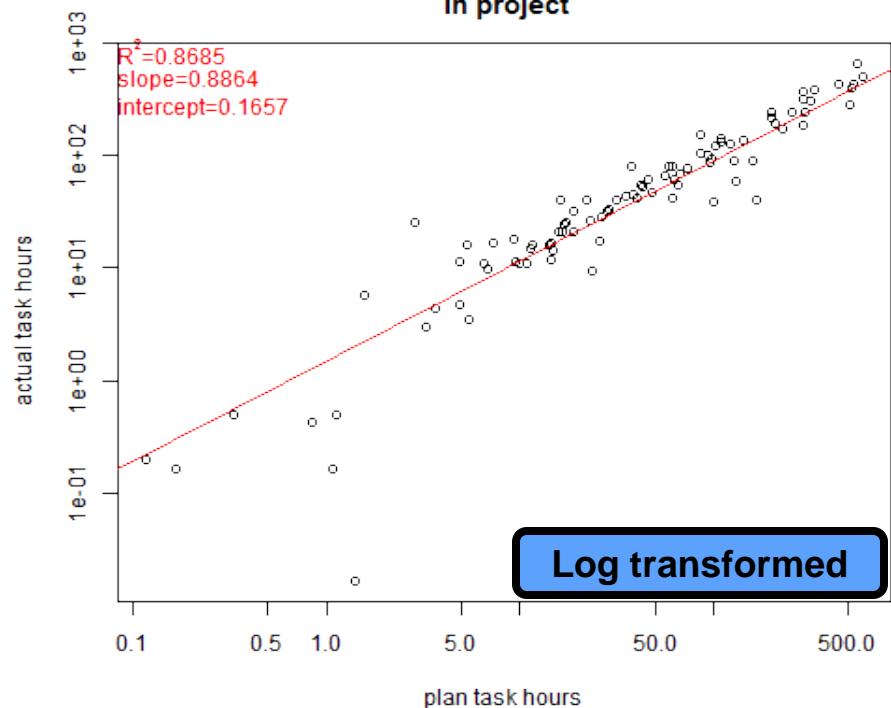
How do Plan and Actual planned project hours compare?

the scatterplot: Plan task hours vs. Actual task hours in project



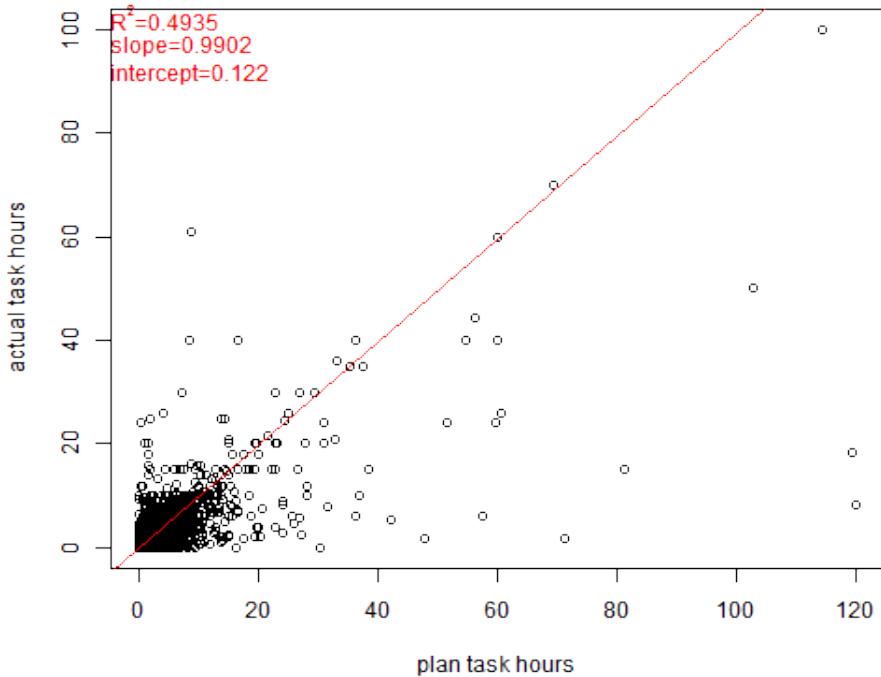
Project level time hours data is high predictable.

the scatterplot: Log plan task hours vs. Log actual task hours in project



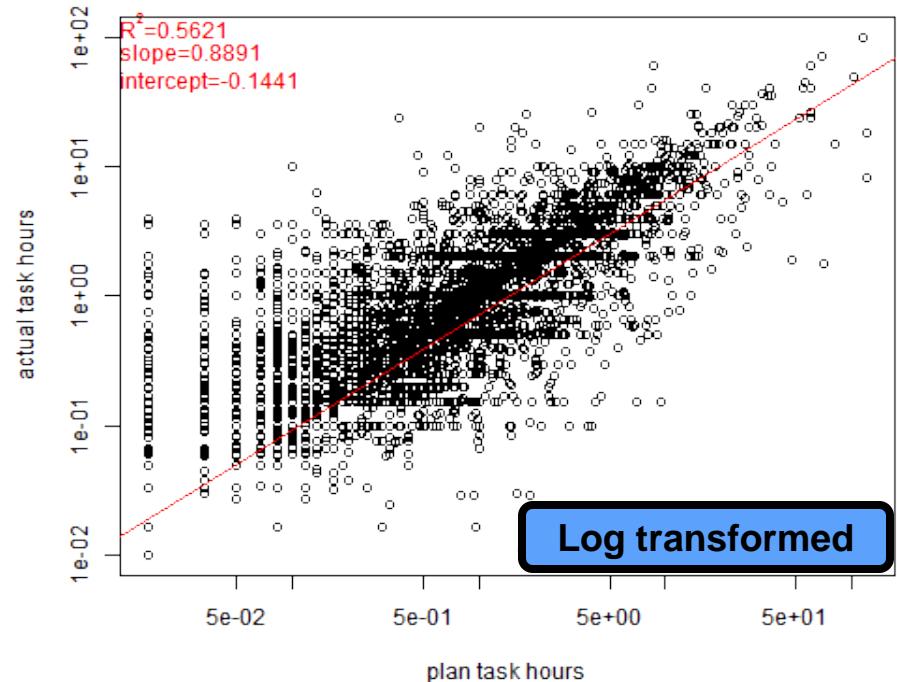
How do Plan and Actual component hours compare? (work item)

the scatterplot: Plan task hours vs. Actual task hours
in component



Log transformed work item level
time hours data is predictable.

the scatterplot: Log plan task hours vs. Log actual task hours
in component



How did they perform to planned schedule?

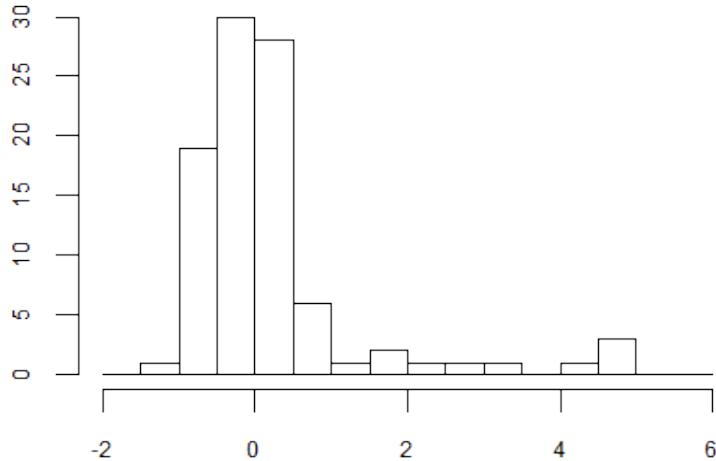
schedule performance =

$$\frac{\text{actual duration} - \text{plan duration}}{\text{plan duration}}$$

Project performance

histogram of schedule performance in all projects
overflow bin = 5

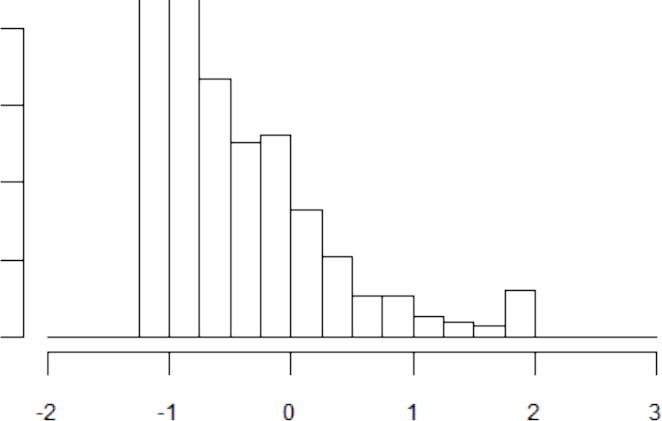
Frequency



Work item performance

histogram of schedule performance in all work items
overflow bin = 2

Frequency



schedule performance



Software Engineering Institute

Carnegie Mellon University

TSP Symposium 2014
William Nichols Nov 4, 2014
© 2014 Carnegie Mellon University

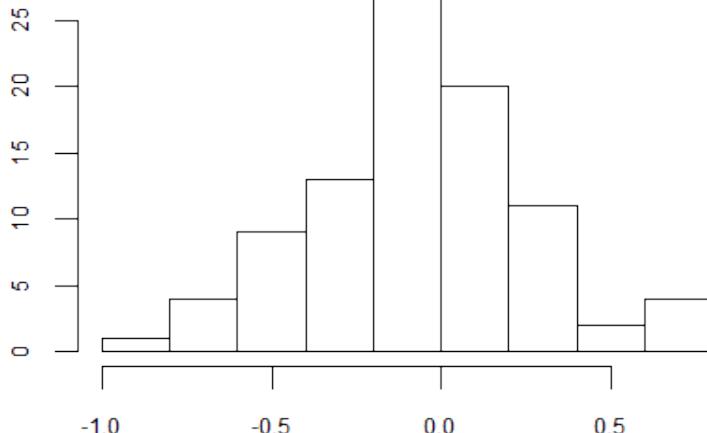
How well did they estimate effort?

effort performance =

Project performance

histogram of effort performance in all projects

Frequency

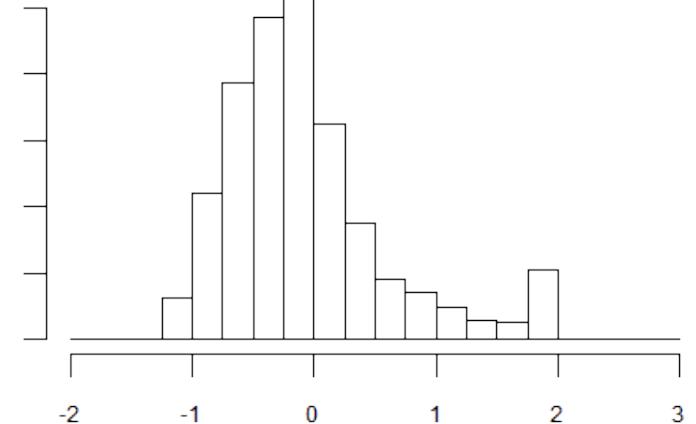


$$\frac{\text{actual effort} - \text{plan effort}}{\text{plan effort}}$$

Work item performance

histogram of effort performance in all work items
overflow bin = 2

Frequency

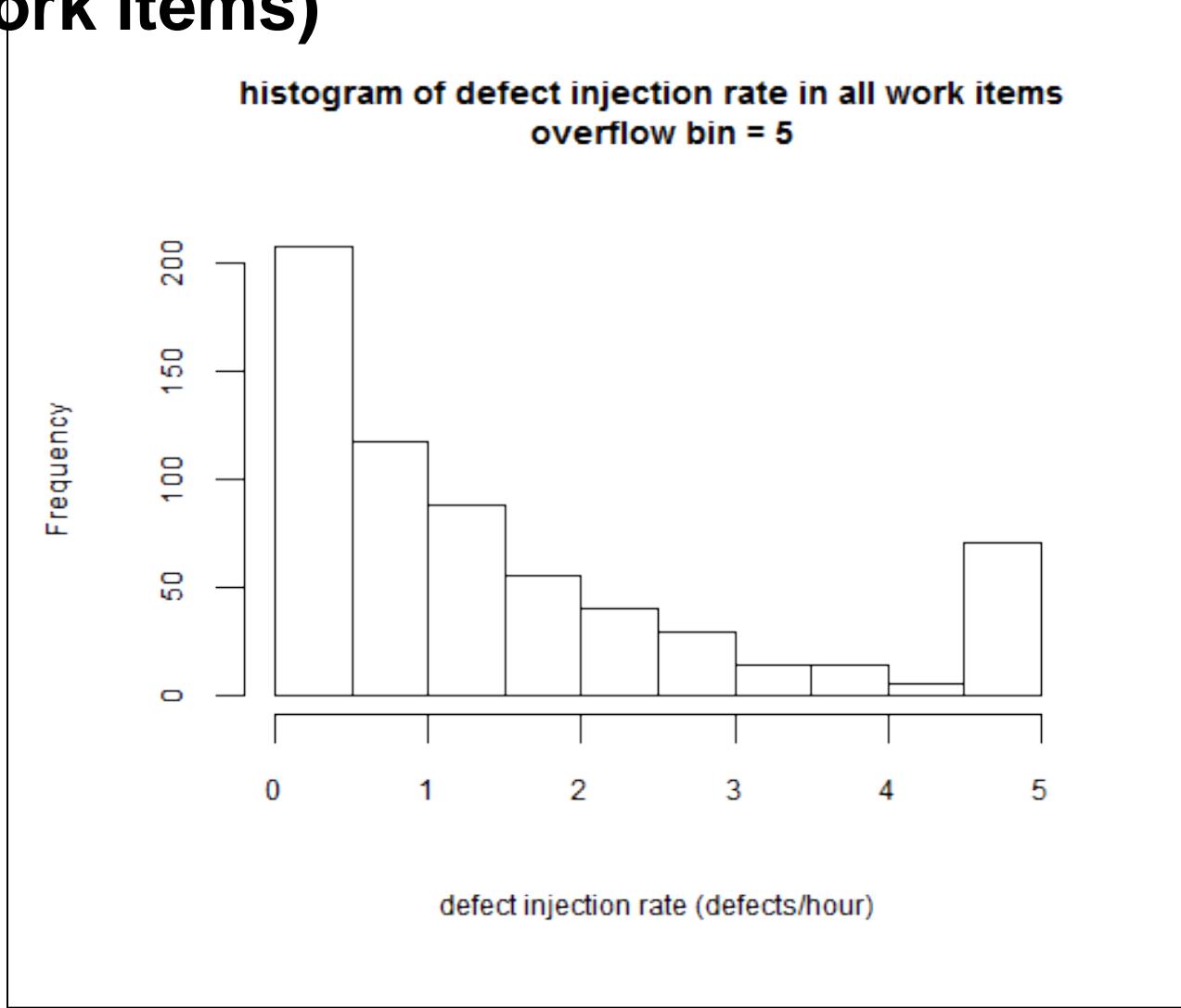


Software Engineering Institute

Carnegie Mellon University

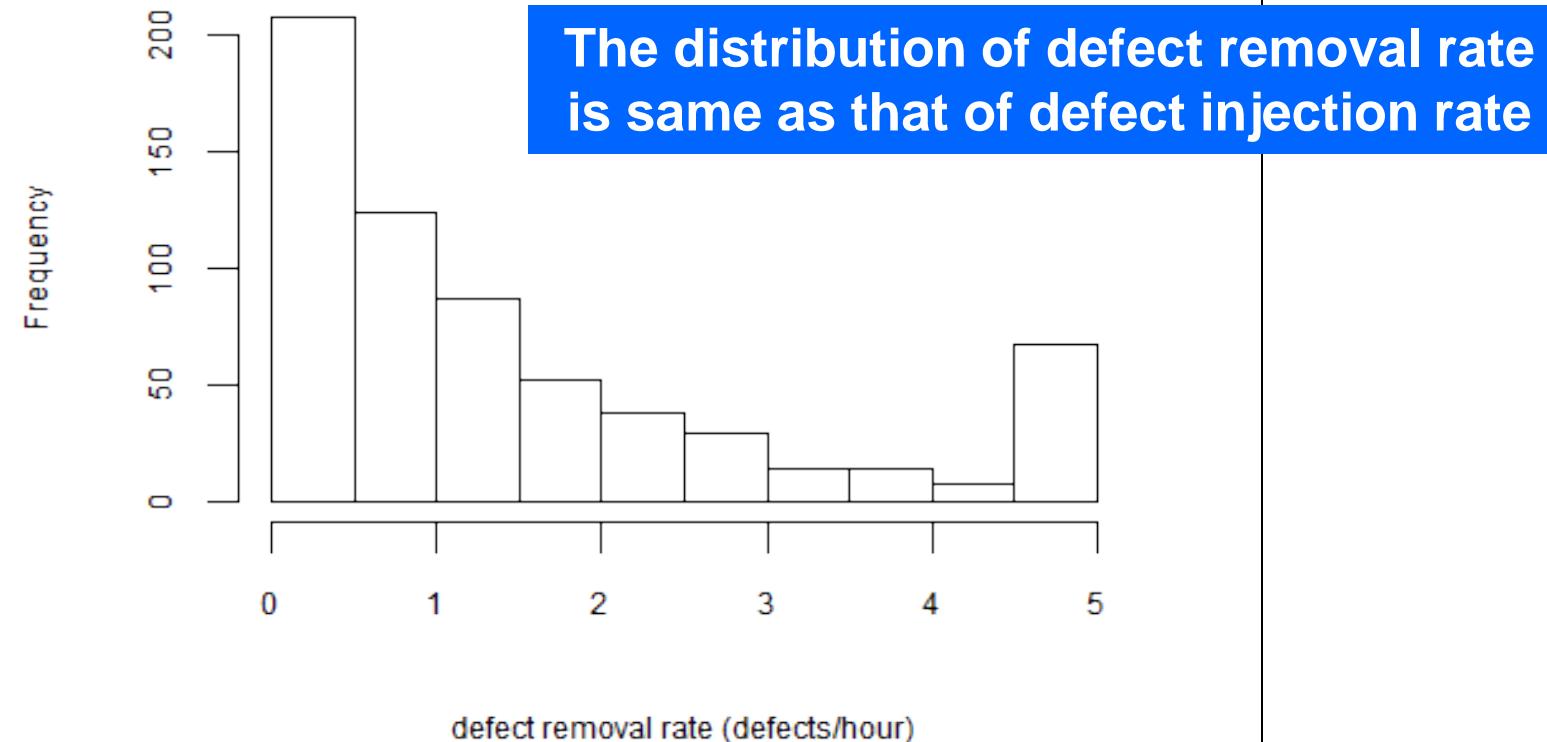
TSP Symposium 2014
William Nichols Nov 4, 2014
© 2014 Carnegie Mellon University

How fast are defects injectioned? (all work items)



What were the defect removal rates? (all work items)

histogram of defect removal rate in all work items
overflow bin = 5

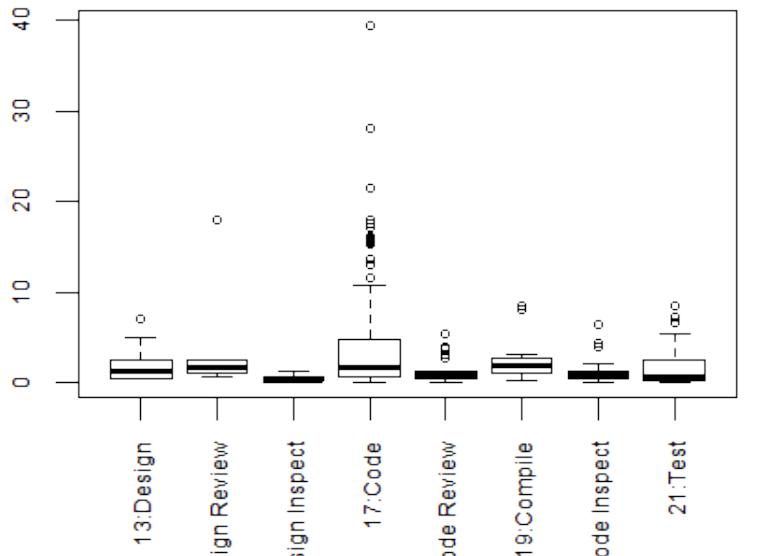


How did defect injection rates differ by phase

All phase

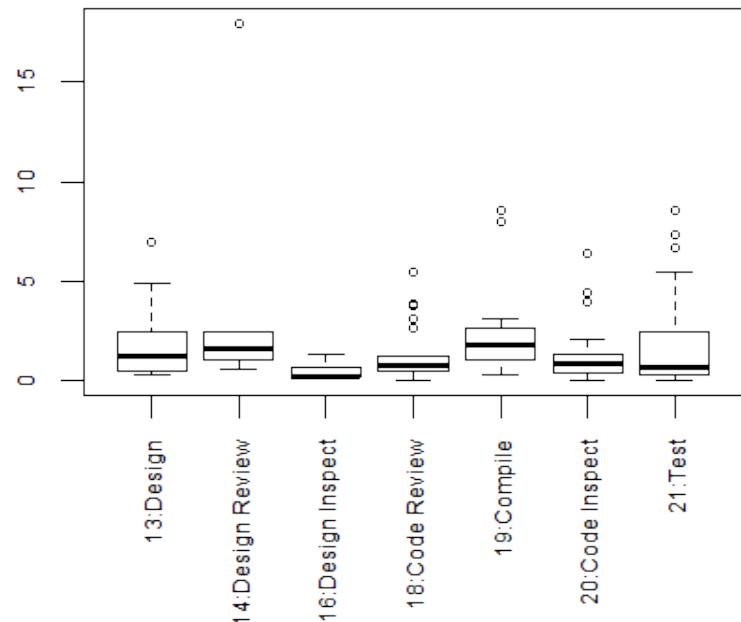
boxplot of defect injection rate in all work items

DIR in code review has wide range and highest median.



Except code phase

boxplot of defect injection rate in all work items



Software Engineering Institute

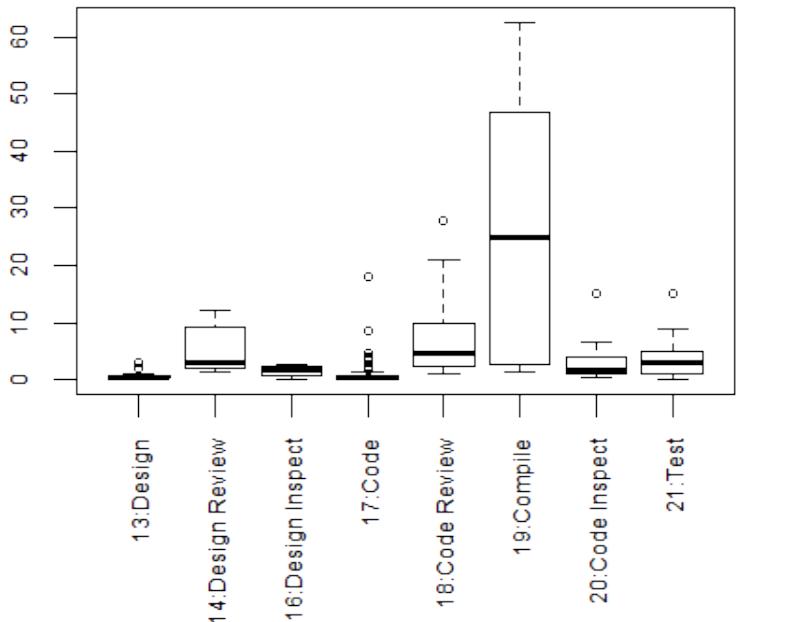
Carnegie Mellon University

TSP Symposium 2014
William Nichols Nov 4, 2014
© 2014 Carnegie Mellon University

How did defect removal rates differ by phase

All phase

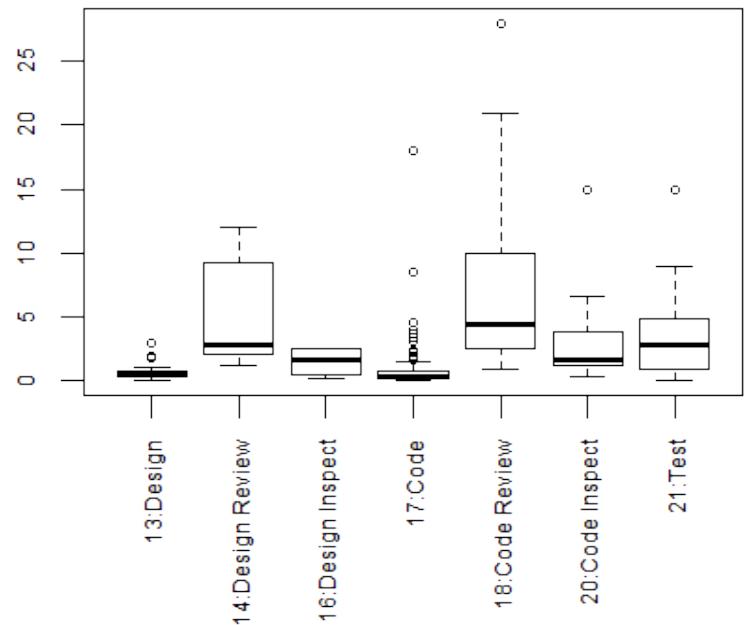
boxplot of defect removal rate in all work items



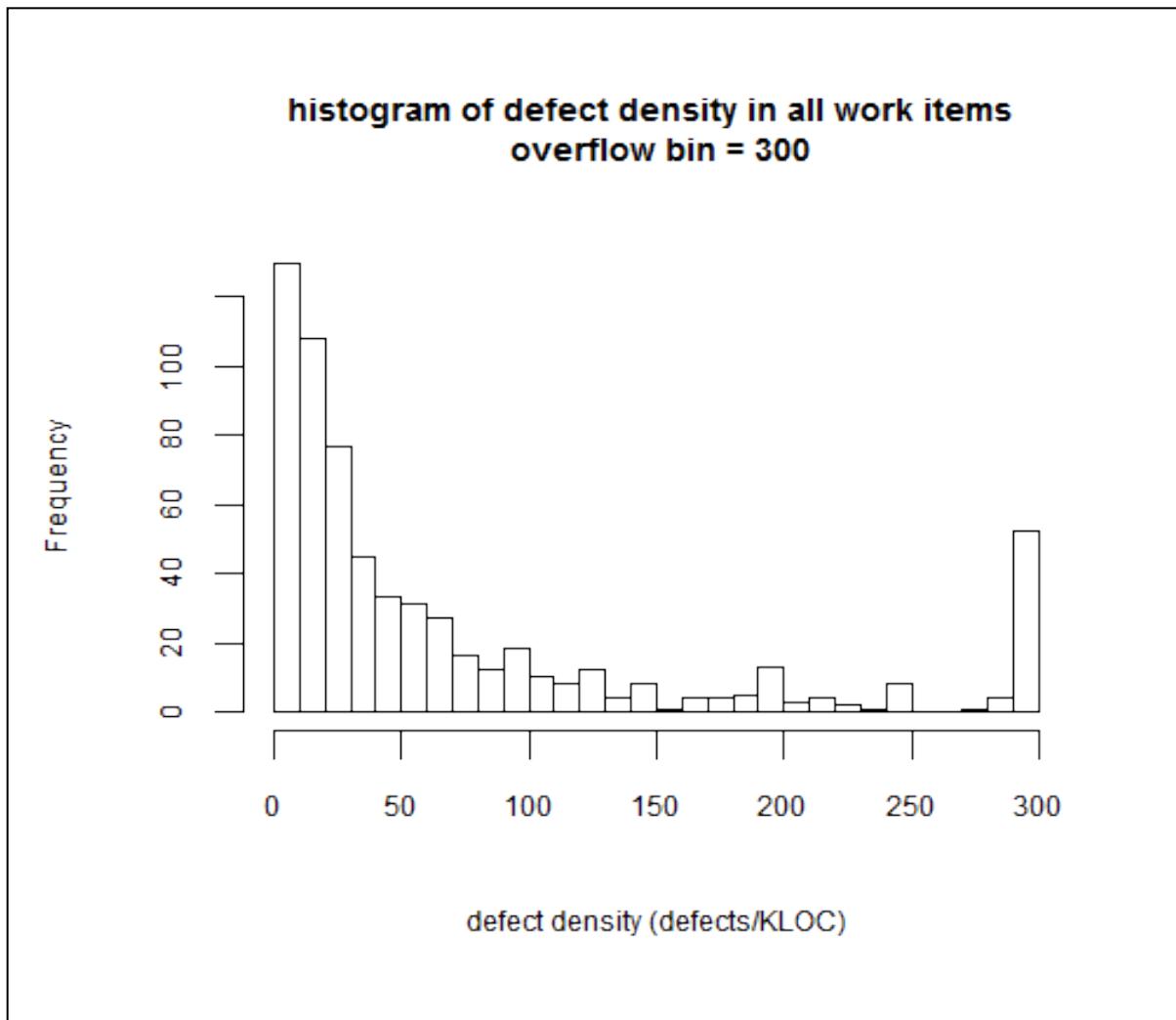
DRR in Compile and DRR in code review are higher than DRR in unit testing.

Except compile phase

boxplot of defect removal rate in all work items



What were the total defect densities



Agenda

1. Introduction
2. SEMPR data analysis
3. Conclusion



Conclusion

SEMPR collects TSP project data for benchmark and analysis

Projects organize in many ways

Benchmarks include

- distributions for defect injection and removal rates
- Ranges of task hours
- Effort estimation accuracy
- Schedule estimation accuracy

Much work remains

- Include more contextual data
- Continue to add projects the database



Acknowledgement

We thank David Tuma of Tuma Solutions for contributing the process dashboard warehouse software.

<http://www.processdash.com/tpdw>



Contact Information

William R. Nichols

TSP/SSD

Telephone: +1 412-268-1727

Email: wrn@sei.cmu.edu

Yasutaka Shirai

Resident Affiliate

Toshiba

Email: yasutaka.shirai@toshiba.co.jp

U.S. Mail

Software Engineering Institute
Customer Relations
4500 Fifth Avenue
Pittsburgh, PA 15213-2612
USA

Customer Relations

Email: info@sei.cmu.edu

Telephone: +1 412-268-5800

SEI Phone: +1 412-268-5800

SEI Fax: +1 412-268-6257

